

实习总结

李伟豪

北京星天地信息科技有限公司

2024/05/28 - 2024/07/31

目录

第一章 实习目标	1
1.1 学习方向及目标	1
1.2 期望	1
第二章 模型的搭建以及学习	2
2.1 Yolo V5	2
2.1.1 模型简介	2
2.1.2 模型环境配置	2
2.1.3 模型训练	2
2.1.4 测试自己制作的数据集	3
2.1.5 实现原理	3
第三章 英文标题 Test	4
3.1 英文标题 Test	4
3.1.1 英文标题 Test	4
参考文献	5
致 谢	6
毕业设计小结	7
附 录	8

第一章 实习目标

1.1 学习方向及目标

学习现在先进的计算机视觉（Computer Vision）以及图形学与 AI 结合的模型如 Neural Radiance Fields（神经辐射场，简称 NeRF），3D Gaussian Splatting(3D 高斯溅射，简称 3dgs)，以及 YOLO（全称 You Only Look Once），同时理解各个模型的实现原理。

1.2 期望

- 完成搭建尽可能多的 AI 模型，配置其环境，并完成训练其预设数据库。
- 学习并理解各个模型的实现原理。
- 自己制作数据并通过基于 AI 的三维重建制作模型。

第二章 模型的搭建以及学习

2.1 Yolo V5

2.1.1 模型简介

Yolo(You Only Look Once) 是一种单阶段目标检测算法，即仅需要“看”一次就可以识别出图片中物体的 class 类别和边界框。Yolov5 是由 Alexey Bochkovskiy 等人在 YOLO 系列算法的基础上进行改进和优化而开发的，使其性能与精度都得到了极大的提升。

2.1.2 模型环境配置

- Python = 3.8.19
- torch = 2.3
- torchvision = 0.18.0
- gitpython = 2.40.1
- opencv-python = 4.9.0.80
- matplotlib = 3.7.5
- pandas = 2.0.3

模型从<https://github.com/ultralytics/yolov5.git> 克隆下来，然后在本地进行配置。

2.1.3 模型训练

模型训练是在 `train.py` 文件中进行的，训练时可以同时输入 `data`¹、`cfg`²和 `weight`³文件；或是 `epochs`⁴和 `batch size`⁵。

如下图：

¹文件中包含了训练集和验证集的路径，以及所有的标注种类

²文件中包含了模型的参数，如学习率、batch size 等

³文件中包含了预训练模型的权重

⁴训练的次数

⁵每次训练的图片数量

```
python train.py --data coco.yaml --epochs 300 --weights '' --cfg yolov5n.yaml --batch-size 128
```

yolov5s	64
yolov5m	40
yolov5l	24
yolov5x	16

图 2-1 YOLOv5 目标检测模型的训练命令

2.1.4 测试自己制作的数据集

从网上找寻了 30 张宝可梦的图片，然后通过 labelImg 工具标注这些图片中比较常见的宝可梦，如皮卡丘、杰尼龟等，然后将这些图片和标注文件放入到一个文件夹中。最后通过 YOLOv5 模型进行训练，得到了一个可以识别这些宝可梦的模型，因为数据集比较小，所以模型的识别率不是很高。

训练结果示例：



图 2-2 YOLOv5 目标检测模型的训练结果

2.1.5 实现原理

YOLOv5 首先会在输入端中将输入图片进行预处理，图像大小调整为模型所需的大小，进行归一化操作，及将像素值缩放到 0 到 1 之间。然后将图片输入到 backbone 网络⁶中，backbone 网络会将图片的特征提取出来，然后将这些特征输入到 neck 网络⁷中，neck 网络会将不同层次的特征进行融合，然后将这些特征输入到 head 网络⁸中，head 网络会将图片中的物体进行预测，得到物体的类别和边界框。

⁶backbone 网络是一个特征提取网络，用于提取图片的特征

⁷neck 网络是一个特征融合网络，用于将不同层次的特征进行融合

⁸head 网络是一个预测网络，用于预测图片中的物体

第三章 英文标题 Test

3.1 英文标题 Test

3.1.1 英文标题 Test

参考文献

致 谢

致谢内容。

毕业设计小结

小结内容。

附 录

附录内容。